View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Harvard University - DASH

DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

Modularity and Intellectual Property Protection

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters.

Citation	Baldwin, Carliss Y., and Joachim Henkel. "Modularity and Intellectual Property Protection". Harvard Business School Working Paper, No. 14-046, December 2013. (Revised January 2014.)
Accessed	February 19, 2015 1:44:32 PM EST
Citable Link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:11591705
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP

(Article begins on next page)

HARVARD | BUSINESS | SCHOOL

Modularity and Intellectual Property Protection

Carliss Y. Baldwin
Joachim Henkel

Working Paper

14-046

January 13, 2014

Copyright © 2014 by Carliss Y. Baldwin and Joachim Henkel

Working papers are in draft form. This working paper is distributed for purposes of comment and discussion only. It may not be reproduced without permission of the copyright holder. Copies of working papers are available from the author.

Modularity and Intellectual Property Protection

Carliss Y. Baldwin[†]
Joachim Henkel*

January 2014

[†] Harvard Business School
cbaldwin@hbs.edu

* Technische Universität München
henkel@wi.tum.de

Copyright © 2014, Carliss Y. Baldwin and Joachim Henkel

Working papers are in draft form. This working paper is distributed for purposes of comment and discussion only. It may not be reproduced without permission of the copyright holder. Copies of working papers are available from the authors.

Modularity and Intellectual Property Protection

Carliss Y. Baldwin,[1] Joachim Henkel[2]

January 2014

Modularity is a means of partitioning technical knowledge about a product or process. When state-sanctioned intellectual property (IP) rights are ineffective or costly to enforce, modularity can be used to hide information and thus protect IP. We investigate the impact of modularity on IP protection by formally modeling the threat of expropriation by agents. The principal has three options to address this threat: trust, licensing, and paying agents to stay loyal. We show how the principal can influence the value of these options by modularizing the system and by hiring clans of agents, thus exploiting relationships among them. Extensions address screening and signaling in hiring, the effects of an imperfect legal system, and social norms of fairness. We illustrate our arguments with examples from practice.

Keywords: Modularity, value appropriation, intellectual property, relational contracts, clans

[1] Harvard Business School, Soldiers Field, Boston, MA 02163, USA, cbaldwin@hbs.edu.

[2] Technische Universität München, Arcisstr. 21, 80333 Munich, Germany, henkel@wi.tum.de.

Copyright © 2014, Carliss Y. Baldwin and Joachim Henkel

Working papers are in draft form. This working paper is distributed for purposes of comment and discussion only. It may not be reproduced without permission of the copyright holder. Copies of working papers are available from the authors.

Modularity and Intellectual Property Protection

INTRODUCTION

Modularity brings many technical and organizational benefits, including the division of labor, reduced cognitive complexity, and higher adaptability and evolvability (Simon, 1962; Garud and Kumaraswamy, 1995; Baldwin and Clark 2000; Schilling, 2000). Yet, despite these well-known technical benefits that support value creation, it is not always straightforward for firms to capture value and protect their intellectual property (IP) in a modular system. In fact, the increased threat to IP has been described explicitly as a drawback of modularity, to be balanced against its various potential benefits (Rivkin, 2000; Pil and Cohen, 2006; Ethiraj, Levinthal, and Roy, 2008).

The risks that modularity poses to IP are illustrated by the history of IBM's System/360, the first "truly modular" computer (Ferguson and Morris, 1993). Peripheral devices such as disk drives, tape drives, and printers could be added as modules to an existing system without difficulty. While customers valued this flexibility, soon after the introduction of System/360, many new firms making peripheral devices entered the market in competition with IBM. Importantly in our context, many of these firms were started by defecting IBM employees (Pugh et al., 1991).

In other cases, however, modularity can help to protect IP, by splitting crucial knowledge into separate modules. Consider the following historical example. In the eighteenth century, Frederick Augustus II, Elector of Saxony, had maintained a monopoly on European porcelain by the simple expedient of imprisoning the inventor in a fortress in Meissen. When the inventor was close to death, Augustus ordered him to divide his knowledge between two successors. One man was told the formula for porcelain paste; the other learned the secrets of making porcelain glaze. Thus, after the inventor died, no one individual could replicate the entire Meissen porcelain-making process (Gleeson, 1998).

In this paper, we will show that modularity can be used to protect IP by enabling companies to disperse and hide information that might otherwise be difficult to protect through the legal system. The relationship between organizational secrets and organization structure has been explored theoretically by

Liebeskind (1996, 1997), Rønde (2001) and Rajan and Zingales (2001). Our analysis builds on this prior work and goes beyond it in the following ways.

First, we distinguish between trustworthy and untrustworthy types of agents. The presence of trustworthy agents in the population makes the option of “doing nothing”—i.e., banking on the possibility that there is no untrustworthy agent among the employees—relevant for the principal, in addition to ex-ante licensing and setting up relational contracts. It also entails a discussion of screening and signaling in order to pick trustworthy agents. Second, we account for relationships among agents by analyzing “clans” in the sense defined by Ouchi (1980) as associations of individuals with common values and beliefs who thus act consistently in most circumstances. Clans may be hired by the principal or emerge through relationship-building within the focal firm. Third, in our analysis of modularity, we allow for different levels of complementarity among modules, and investigate how clans and modularity interact in determining the preferred organizational structure. Lastly, when clans or modules are asymmetric, our model—based on rational choice theory—indicates that members of larger clans and agents working on less valuable modules should be paid less. However, we argue, this strictly rational approach may be problematic when agents are socialized within relationships to value reciprocity and fairness.

Our main results are the following. For the base case of a one-module system, “doing nothing” is the best option for protecting the principal’s IP if the percentage of trustworthy agents in the population is high. Ex-ante licensing is optimal if the percentage of trustworthy agents is low, and the intensity of competition in duopoly is also low (so that the license is valuable). Establishing relational contracts with agents (i.e., paying them above their efficiency wage not to defect) dominates both doing nothing and ex ante licensing if the intensity of competition is high, the number of agents is low, and the percentage of trustworthy agents is below a threshold. Finally, a decline in the number of agents needed to perform the focal tasks increases the range of parameter values for which doing nothing or a relational contract is preferred to licensing.

Hiring clans of agents, or promoting their emergence among employees, is a way for the principal to mitigate the risk of expropriation. Assuming that all members of a clan act together, the presence of clans in effect reduces the number of independently acting agents among the employees. This makes both doing nothing and a relational contract more attractive relative to licensing.

Modularizing the focal technical system has similar effects to hiring clans. Each module has fewer agents than the whole system and is worth less, hence doing nothing and relational contracts increase in value relative to licensing. This effect increases as modules exhibit higher levels of complementarity. In this context, we carve out the important distinction between modularity in use and modularity in production and design. While the former facilitates imitation and substitution (Rivkin, 2000; Pil and Cohen, 2006; Ethiraj, Levinthal, and Roy, 2008), the latter mitigates the risk of imitation by agents.

Clans and modularity interact in an important way. As long as all members of a clan work on the same module, their protective effects against expropriation reinforce each other. But if members of a clan are spread across modules and can share their knowledge, then clan members will have access to knowledge that module boundaries could have kept hidden from them. In that case, clans partly defeat the protective effect of modularity.

When clans or modules are asymmetric, our model—based on rational choice theory—indicates that members of larger clans and agents working on less valuable modules should be paid less. Also cohorts of new hires that, due to improved screening or signaling technology, have a higher share of trustworthy agents may receive lower payments. We shall argue, however, that the use of rational choice theory to predict agents' behavior may not be appropriate if agents value relationships characterized by reciprocity and fairness.

Our paper is organized as follows. In the next section, we review the relevant literature. We then begin our formal analysis by introducing and analyzing the base case of a one-module system. We analyze the impact of clans and go on to study the impact of modularity and its interaction with clans. After discussing several extensions of our analysis, we provide illustrative examples from practice. We conclude the paper by describing the limitations of our analysis, implications for scholars and managers, and directions for future work.

BACKGROUND

Knowledge may be a source of profits and competitive advantage, so long as it cannot be expropriated, imitated, or substituted (Teece, 1986, 2000; Barney, 1991). IP rights may protect against expropriation and imitation, but vary in strength by jurisdiction and industry (Maskus, 2000; Zhao, 2006; Kyle and McGahan, 2009; Branstetter et al., 2011). When formal IP rights are weak, relational contracts may afford protection against expropriation. As we will show, they may be particularly effective in conjunction with modularity. In this section, we review the relevant strands of literature.

Relational Contracts in Economics, Law and Sociology

The economic theory of the firm is concerned with the location of boundaries between companies (Coase, 1937). Grossman and Hart (1986) and Hart and Moore (1990) developed a theory of the firm based on optimal allocation of property rights. Brynjolfsson (1994) applied their reasoning to knowledge and intellectual property. We follow Brynjolfsson in focusing on knowledge as an asset, and we follow Hart and Moore (1990) in defining “property” as the ability to exclude others from using the asset. We differ from these prior works, however, in that we do not consider property rights to be secure.

Baker, Gibbons and Murphy (2002) extended Grossman, Hart, and Moore’s theoretical framework to include so-called “relational contracts.” In a relational contract, deviations from cooperative behavior can be punished by terminating the relationship. As long as the reward to deviation is less than the continuation value of the relationship, parties to the contract will cooperate without state enforcement (Kreps, 1990; Greif, 1998; Gibbons and Henderson, 2012). Relational contracts are thus self-enforcing (Telser, 1980; Baldwin, 1983; Greif, 1998). They can be modeled as repeated games (Bull, 1987; Kreps, 1990; Baker et al., 2002). In practice, they take different forms including unilateral contractual payments, bilateral contractual payments and equity-based alliances (Oxley, 1997).

It should be noted that the concept of relational contract in economics differs from that in law and sociology. In law, for example, Macneil (1978, 1985) defines relational contracts as associations that have significant duration and involve close personal relationships, with “entangling strings of friendship, reputation, interdependence, morality, and altruistic desires” (Macneil, 1987, p. 276). In contrast, in economics and game theory, agents are assumed to be purely calculative about a continuing relationship, that is, they constantly ask the question “is it worthwhile for me to stay in this relationship or not?” (Williamson, 1993). In what follows, we use the term “relational contract” as in economics and game theory, to mean a self-enforcing agreement between self-interested, value-maximizing agents.

However, the difference in perspectives leads to different interpretations of the concept of “trust” which is a central focus of our analysis. Scholars in sociology, for example, Granovetter (1985), define trust as the expectation of non-calculative, benign action by another agent, and see it arising from a combination of embeddedness in social networks and repeated personal interaction (Uzzi, 1997). Dyer and Singh (1998) argue that informal safeguards based on trust generate greater “relational rents” than formal safeguards, but empirical support of this contention is mixed at best (Sako, 1998). Taking the economic perspective, Williamson (1993, pp. 475-479) acknowledges that trustworthy behavior in commercial relationships can be elicited by environmental conditions, including embeddedness in social networks and cultural norms and expectations. But, he argues, such actions flow from agents’ calculations of long-term self-interest. Supporting this point of view, there is evidence that transaction hazards (i.e., temptations) increase the probability of untrustworthy behavior (Poppo, Zhou and Zenger, 2008). But there is also countervailing evidence from psychology that cost-benefit analysis does not affect dishonest behavior (Mazar, Amir and Ariely, 2008).

In general it is impossible to infer from actions alone whether a given agent’s trustworthy behavior is motivated by calculations of long-term interest, concerns about social sanctions, personal integrity, or a combination of these factors (Posner, 2002; Poppo and Zenger, 2002). However, there is observable variation in behavior across individuals and populations. Rotter (1980) presents evidence that some individuals are innately more trusting and trustworthy than others. And the probability that trustworthy behavior arises from social norms and moral beliefs vs. direct payments varies across cultures and is sensitive to surrounding institutions (Yamagishi and Yamagishi, 1994; Meier, 2006; Bjørnskov, 2007). Still one can never be 100% sure that the person one is dealing with is trustworthy—trust can always be abused (Granovetter, 1985).

In our model, we distinguish between agents, called “trustworthy,” for whom social norms and moral beliefs are sufficient to prevent defection and agents, called “untrustworthy,” who require financial compensation not to defect. We allow the percentage of trustworthy and untrustworthy agents to vary across populations, and show how the focal firm will condition its strategy on this variable. We admit that this is a crude way to capture the subtle nuances of relationships in organizations, but we believe that deciding whether to trust (vs. pay) one’s agents is a true strategic choice for firms whose competitive advantage rests on protecting organizational secrets.

Organizational Secrets and the Problem of Expropriation

In a seminal paper, Liebeskind (1997) opened up the topic of protecting organizational secrets by discussing the benefits and costs of keeping a firm’s unique knowledge safe from public view. She framed secrecy as an economic tradeoff, and discussed various methods used by firms to protect their secrets. Rønde (2001) then constructed a formal model in which a principal, who needs to grant agents access to his knowledge in order to commercialize it, fears that they will expropriate it. (Rønde’s agents are strictly calculative.) The principal can either grant all agents full access, or divide the task at hand and provide to each agent only the information she needs for her task.

Rajan and Zingales (2001) analyze how a principal can contain the risk of knowledge expropriation through the firm’s hierarchical structure and size. In a vertical hierarchy, agents on lower levels are assumed to have limited access to the principal’s knowledge due to their larger distance from the knowledge source and specialization to their direct superiors. This specialization and the resulting loyalty drive the result that physical-capital-intensive industries should be characterized by steep hierarchies, while in human-capital-intensive industries flat hierarchies should prevail.

Finally, the idea of dividing knowledge in order to capture more of its value figures in Anton and Yao's (2005) model of a sale of IP subject to Arrow's (1962) information paradox. They suggest splitting the knowledge in such a way that one part is protectable and informative about the value of the IP overall, and selling this part first. They find this approach to be more profitable for the seller than a bundled sale of both parts.

Clans

One possibility for the principal to protect his knowledge from expropriation is to employ agents who defect, or stay on, in groups. Ouchi (1980) defines a "clan" as an association of individuals who have been socialized to have common values and beliefs and thus act consistently in most circumstances. Clan members who deviate may also be punished in some fashion, for example by loss of access to the clan, ostracism, or shunning. A clan "resembles a kinship network but may not include blood relations" (Ouchi, 1980: 134).

The focus of Ouchi's (1980) and most subsequent work in this context is on clan control of a firm as an alternative to market or bureaucratic control mechanisms. The firm's employees, effectively, constitute the clan. However, the clan may also be a subset of the firm's employees. In that case, the clan's goals can be incongruent to those of the firm, as for example in Johnson et al.'s (2002) study of international joint ventures or Groysberg and Abrahams' (2006) analysis of "liftouts" (defined as the hiring away of whole organizational units or teams). Indeed, teams whose members work closely together may become embedded in their own social network and thus develop the characteristics of a clan—more precisely, a "social-integrative clan" in the nomenclature of Alvesson and Lindkvist (1993). Alternatively, the principal may hire clans defined by e.g. nationality or family ties.

As we will show, employing clans can both aggravate and mitigate the threat of expropriation of knowledge. As a means of protection it is most effective in conjunction with modularity.

Modularity

According to the theory of modularity, firms can divide complex technical systems into components ("modules") that can be designed independently but function together as a whole. Three key concepts are worth noting. First, the modular structure of a technical system is a choice that system architects make (von Hippel, 1990; Ulrich and Eppinger, 1994; Baldwin and Clark, 2000). Most complex technical systems can be designed to be more or less modular, and the boundaries between modules can be located in different places (Mead and Conway, 1980; Ulrich and Eppinger, 1994). Second, if the separation of modules is done properly, the design decisions taken with respect to one module will not affect decisions taken in other modules. Design tasks can then be allocated to different organizational units or firms (von Hippel, 1990; Langlois and Robertson, 1992; Sanchez and Mahoney, 1996). Third, just as modules can be separated in terms of their underlying design decisions, knowledge about modules can likewise be separated. As long as they can access the design rules specifying the interfaces, Module A's designers do not need to have specific knowledge about Module B's internal structure. Thus the designers of each module have (potentially) exclusive knowledge. Conversely, designers working within a module must share knowledge or risk jeopardizing the success of their efforts.

While the technological and organizational consequences of modularity have received a great deal of scholarly attention, the strategic consequences—i.e., how modularity affects competition among firms—have been studied less widely. Rivkin (2000), Pil and Cohen (2006) and Ethiraj, Levinthal, and Roy (2008) argue that modularity poses a strategic trade-off for innovators. On the positive side, it allows the focal firm to innovate faster and thus stay ahead of would-be imitators. Further advantages, given inter-firm compatibility, may be the chance to mix and match modules from different sellers (Matutes and Regibeau, 1988) and to upgrade individual modules selectively (Garud and Kumaraswamy, 1995). An innovator may even invite competitive entry through modularity in order to promote a market segment as a whole (Garud and Kumaraswamy, 1993). On the negative side, modularity makes a firm's products easier to imitate (Rivkin, 2000; Pil and Cohen, 2006; Ethiraj, Levinthal, and Roy, 2008). In this paper, we add to this prior work by looking at how modularity affects the threat of expropriation of IP, and carve out the important distinction between modularity in production (which partitions and thus can protect organizational knowledge) and modularity in use (which exposes the firm to external competition on modules).

THE BASE CASE: PROTECTING KNOWLEDGE IN ONE-MODULE SYSTEMS

When someone possesses knowledge and wants to realize its value, he must generally employ individuals and contract with suppliers who will turn that knowledge into a working product or process. But in doing so, the principal must (almost always) reveal that information to those agents, subject to the modular division of the system. Those agents, in turn, could set up a rival establishment or reveal the knowledge to competitors (for clarity of exposition, we focus on the first possibility). This threat is well-known in law and economics and has been discussed by Teece (1986), Liebeskind (1997), Rønde (2001), and Rajan and Zingales (2001). Organizational knowledge may be protected via trade secrets law and non-disclosure agreements, but such protection is imperfect and its effectiveness varies by jurisdiction (Sherwood 1990; Oxley, 1999; Fisk, 2001; Lemley 2008; Marx 2011).

Model set-up

We first consider the simplest case: a one-module system, in which each design decision is related to all others. Thus, people working on the module must have unrestricted access to all relevant knowledge in order to address the system's interdependencies. This leaves the principal vulnerable especially when, as we assume, property rights or contracts over knowledge are not enforceable within the governing legal system.

Let the total number of agents who need access to the principal's knowledge be denoted by N. The agents fall into two types. The first type, called "trustworthy," will under no circumstances defect. The second type, called "untrustworthy," will defect if it is in their economic interest to do so. Each agent knows his or her own type, but not the types of the other agents. The probability, t, that any given agent is trustworthy is known to both the principal and all agents. We model t as exogenous for the time being. We assume that untrustworthy agents decide independently whether to defect or not.[3] Apart from not knowing the other agents' types, all agents have full information about the parameters and the structure of the game.[4]

To keep the analysis tractable and focused on our aim of studying the impact of relationships and modularity, we assume that only two firms can profitably operate in the market.[5] We define V as the value of the monopoly, and αV as the net value, per firm, of duopoly. A defecting agent who successfully sets up a rival firm appropriates αV.[6] Finally, we assume that $0 < 2\alpha V < V$; otherwise an efficient principal would want to create a second establishment of his own accord. In general, V can be ex ante uncertain, in which case α corresponds to a proportionate equity share in the monopoly. There may be reasons to prefer one form of relational contract (e.g. an equity alliance) to others, but in general, those concerns lie outside the scope of our model (cf. Oxley, 1997, 1999 for an analysis of different contract forms).

The principal's options

In dealing with the risk of defection the principal has three options: to do nothing and bank on the possibility that all his agents are trustworthy; to license his technology to the highest bidder before hiring agents; and to enter into relational contracts with his agents. We analyze each in turn and then determine

---

[3] The timing of moves is as follows. Each period is divided into two sub-periods. In the first sub-period, agents simultaneously and independently decide whether to defect and the defectors leave. In the second sub-period, the principal learns if any have defected and pays the agents accordingly. The defectors, if any, collect and split their reward. Then, conditional on no defections, the game is repeated. There is no last period of the game, although it may end probabilistically as a result of exogenous events.

[4] In his analysis of social norms, Posner (2002) makes similar assumptions about the existence of trustworthy (cooperative) and untrustworthy (uncooperative) agents. He then considers how cooperative agents use conformance with social norms to credibly signal their type. Below we address the principal's and agents' incentives to invest in better screening or signaling technology, which would help them to increase t among the principal's employees. For now, though, we treat t as exogenous.

[5] The general case is extremely complicated because the payments necessary to keep untrustworthy agents loyal are determined recursively. Depending on the exogenous value of a k-firm oligopoly vs. a $k+1$-firm oligopoly, and the fixed cost of setting up a new firm, an intricate sequence of endogenous agent payments can be found through backward induction. The sequence is not necessarily monotonic.

[6] Also when joining an existing competitor, the defecting agent may be able to appropriate the entire value of αV, e.g., if several existing firms compete for her knowledge.

the principal's best option. For simplicity, we assume all parties are risk neutral, although this assumption is not essential to the results.

Doing nothing, licensing

If the principal "does nothing," then his monopoly is preserved if and only if all N of his agents are trustworthy, which happens with probability t^N. His expected payoff then is

$$\Pi_{\text{nothing}} = [t^N + \alpha(1 - t^N)]V \ . \tag{1}$$

Note that the value of doing nothing declines as N, the number of agents goes up. In effect, each additional agent "in the know" increases the probability that one of them will be untrustworthy.

Alternatively, the principal can, before hiring agents, forgo his monopoly, and license his technology to the highest bidder who will then set up a second competing establishment. By our simplifying assumption that only two firms can profitably operate in the market, the existence of the second establishment makes defection unattractive to agents at both establishments. Assuming more than one bidder and Bertrand competition, we obtain as the value of this option:

$$\Pi_{\text{licensing}} = 2\alpha V \ . \tag{2}$$

Relational contracts

As a third option, the principal can set up a self-enforcing relational contract with the agents. Following common practice in economics, we model a relational contract between a principal and his agents as a repeated game in which the principal essentially pays the agents not to defect (Bull 1987; Baker et al. 2002).

To set up a relational contract with calculative agents, the principal promises to pay each agent a bonus above the competitive wage with a present value of ζV if nobody defects and zero otherwise.[7] The minimum bonus is affected by the principal's need to make the contract self-enforcing. Specifically, if $\zeta < \alpha$ then loyalty by all (untrustworthy) agents cannot be an equilibrium since each can do better by defecting.

To fully specify the game, we must describe what the untrustworthy agents expect to happen when two or more defect. One possibility is that each defector immediately builds a new establishment. Since by assumption, only two establishments profitably operate, the defectors will all incur losses. The game among agents is essentially a game of "chicken," and the unique Nash equilibrium (if $\zeta < \alpha$) is for one and only one agent to defect. Alternatively, potential defectors might expect to come together and split the value of the second establishment amongst themselves. With $\zeta < \alpha$, all untrustworthy agents will defect (since agent payments for those who stay will go down to zero after defection of one or more others). This game is essentially a prisoner's dilemma (or social dilemma). Each agent gains at the margin by defecting, but in aggregate the defectors are worse off than if they had been loyal.

Interestingly, from the principal's standpoint the design of the relational contract does not depend on the agents' conjectures about the behavior of other agents. Whether the game is chicken or a prisoner's dilemma, if $\zeta < \alpha$, then each Nash equilibrium in pure strategies is characterized by one or more untrustworthy agents defecting. And one defection suffices to end the principal's monopoly.

Thus, to bring about an "all stay" equilibrium, the principal must set $\zeta = \alpha$, paying every untrustworthy agent an amount whose value is equal to the total defection reward, αV. And because (by assumption) the principal cannot distinguish between untrustworthy and trustworthy agents, *all* agents

[7] To keep notation simple, we assume that agents live forever. Assuming, instead, a constant probability of dying in each period would keep our results qualitatively unchanged.

must receive a stream of payments whose value equals αV. Thus the total cost of protecting the principal's knowledge against unauthorized use by agents is $N\alpha V$ and the value of this option is:

$$\Pi_{\text{payments}} = [1 - \zeta N]V = [1 - \alpha N]V \; . \tag{3}$$

The principal's best option

We start by comparing "doing nothing" to licensing. It is straightforward to show that for low values of t (an untrustworthy population), the principal will choose licensing, while for high values of t, he will hope to preserve the monopoly and do nothing. The following proposition provides more detailed results. All proofs are in the Appendix.

> ***Proposition 1 (a)*** *For a one-module system, if (1) property rights and contracts are not enforceable; (2) the principal cannot distinguish between trustworthy and untrustworthy agents; and (3) the defection reward (equivalent to the licensing payment) is* αV *, then the principal should opt to license his knowledge if the percentage of trustworthy agents:*
>
> $$t < t^{\dagger} = \left(\frac{\alpha}{1-\alpha}\right)^{1/N} \; ; \tag{4}$$
>
> *The principal should do nothing if* $t > t^{\dagger}$*, and is indifferent if* $t = t^{\dagger}$*.*
>
> ***(b)*** *The threshold value* $t^{\dagger}$ *increases in both* α *and N.*

We now compare the payoff obtainable using relational contracts (Equation 3) to those from doing nothing and licensing (Equations 1 and 2 respectively). This leads to:

> ***Proposition 2 (a)****. Under the same assumptions as Proposition 1, the principal can achieve an "all-stay" equilibrium in a relational contract by paying each agent an annuity whose present value, denoted* ζV *, equals the total defection reward* αV *.*
>
> ***(b)*** *Setting up a relational contract is the best policy for the principal if two conditions hold:*
>
> $$N < \alpha^{-1} - 2 \; ; \quad \text{and} \quad t < t^{*} = \left[1 - \frac{N\alpha}{1-\alpha}\right]^{1/N} \tag{5}$$
>
> *If either condition is violated, then one of the other options (do nothing or licensing) is preferable.*
>
> ***(c)*** *The threshold value* t^{*} *decreases in both* α *and N.*

To see how parameter changes affect the principal's best option, we calculated for each pair of options under which conditions one is equal or superior to the other. Solving the resulting conditions for α, we obtain:

$$\begin{aligned} \Pi_{\text{nothing}} \geq \Pi_{\text{payments}} &\iff \alpha \geq \frac{1-t^N}{N+1-t^N} \\ \Pi_{\text{nothing}} \geq \Pi_{\text{licensing}} &\iff \alpha \leq \frac{t^N}{1+t^N} \\ \Pi_{\text{payments}} \geq \Pi_{\text{licensing}} &\iff \alpha \leq \frac{1}{N+2} \end{aligned} \tag{6}$$

Figure 1 shows the (t,α) parameter space for N=1 and N=5.[8] We can divide the parameter space into three regions defined by the principal's respective best option. In Region L, licensing is optimal; in Region N, doing nothing; and in Region P, paying agents. The subscript of each label indicates which option is second-best. For continuity reasons, the first-best option in a given sub-region is second-best in the neighboring sub-region. For example in sub-region L_P, licensing is first-best and payments are second-best, while in the adjacent sub-region P_L, payments are first-best and licensing is second-best.

We now establish a basic result that will be useful when we incorporate clan relationships and modularity into the analysis. From Equations 1, 2, and 3, it is clear that reducing N, the number of agents, increases the value of doing nothing and agent payments, while having no effect on the value of licensing. Since the value of each option increases weakly as N decreases, the maximum of these values also increases weakly. We summarize:

> ***Proposition 3.*** *Other things equal, a decline in the number of agents N is either beneficial or value-neutral for the principal. With declining N, the region in (t,α) parameter space in which licensing is optimal shrinks, while the other two regions expand.*

In the following sections, we show that employing clans and modularizing the product architecture each have the same effect as reducing N.

---

[8] Note that, even in large firms, the relevant number N of agents with access to the principal's knowledge may be quite low.

Relationships among agents: Clans

By definition, members of the same clan act together. We can thus assume that the clan follows its leader (or, if it does not have one specific leader, behaves as if it had one), and so the probability that any given clan is trustworthy equals the probability t that its leader is trustworthy. The definition implies that individuals in clans can coordinate their actions.

Suppose the N agents are divided into L "clans" of size N/L, $1 \leq L \leq N$. Necessarily, N and L are both integers: the fractional parts of N/L can be interpreted as agents who work part-time. $L=N$ is the condition where each agent acts as an individual, $L=1$ is the condition where all agents belong to the same clan. We assume that defecting clans will split the reward to defection equally amongst all members of the clan[9], and restrict the analysis for simplicity to clans of equal size (we will address the asymmetric case below in our discussion of fairness).

Under these assumptions, employing L clans is mathematically equivalent to employing L agents, while keeping parameters t and α the same. Each clan behaves as a single decision-maker and each can expect a total defection reward equal to αV. Thus, the value of doing nothing becomes: $\Pi_{\text{nothing, clans}} = [t^L - \alpha(1 - t^L)] \cdot V$; the value of licensing is unchanged; and the value of the agent payment strategy becomes: $\Pi_{\text{payment, clans}} = [1 - \alpha L] \cdot V$.

We can now apply Proposition 3. Since $L < N$, employing clans makes both agent payments and doing nothing more attractive relative to licensing. In the (t, α) parameter space, Regions N and P expand while Region L must shrink. If the principal starts and ends the day in Region L, then employing clans is value-neutral. In contrast, if, with clans, the principal ends the day in Region N or Region P, then he is unambiguously better off employing clans vs. employing individualistic agents.

[9] If the clan is hierarchical, such that higher-ranking members are paid more than lower-ranking members, we assume the payments will be similarly apportioned, if the clan defects or does not defect.

THE IMPACT OF MODULARITY

The option to hire agents linked by clanship may or may not be available to the principal. However, he has full control over the design of the system subject to technological constraints. In particular, he can design the system in a modular fashion and thus divide the relevant knowledge into separate modules. Importantly, we assume that the principal either sells only complete systems rather than individual modules (Fixson and Park, 2008), or is able to protect the system's interfaces such that no third party products can be attached to the principal's system.

Modeling modules and complementarity

To start with, consider the simplest case: a symmetric split of the system into M modules, each worth V/M. Assume that, as with the entire system, only one competing establishment can be profitably set up for each module. A defecting agent secures the reward $\alpha V/M$, which is also the principal's duopoly payoff from that module. In this situation, the principal's payoffs under the three strategies are given by Equations 1 – 3 above, with N replaced by N/M. Also, V is replaced by V/M and the entire equation multiplied by M: the latter two Ms cancel each other out, leaving only N/M in place of N. Thus, this type of modularization has the same effect as reducing the number of agents. By Proposition 3, it is beneficial or neutral for the principal.

We now relax the assumption that the overall value of the system is the additive sum of the modules' value. To simplify the analysis and reduce the vast number of combinatorial possibilities, we assume there are two types of defection rewards: a module reward that can be claimed by defectors of any module, and a system reward that can only be claimed if there is a defector *in each module*. We further assume that the total defection reward for a system of M modules can be expressed as a convex linear combination of the two types:

$$\text{Total Defection Reward} \equiv (1-\beta)\left(\sum_{i=1}^{M}\delta_i\right)\frac{\alpha V}{M} + \beta\left(\prod_{i=1}^{M}\delta_i\right)\alpha V \quad , \tag{7}$$

where β is a number between 0 and 1, and δ_i equals 1 if at least one agent from module *i* defected, otherwise 0.[10] Thus module defectors can claim a reward equal to $(1-\beta)\alpha V/M$ for their own module, and in addition, if every module has a defector, the defectors as a group can claim the system reward equal to $\beta\alpha V$. Notice that, with M=1, the total defection reward equals αV even with only one defector. Hence the one-module system previously analyzed is subsumed in this broader specification.

In Equation 7, β is a parameter that allows us to "tune" the degree to which modules are complementary. If $\beta = 0$, then each module has a separate stand-alone value $\alpha V/M$, and the defectors' total reward is simply the sum of module rewards. In contrast, if $\beta = 1$, the modules are strict complements, and all must be present for the defectors to realize any reward at all. For β between zero and one, the modules have some stand-alone value, but there is additional value derived from putting all the pieces together.

Doing nothing, licensing

We start, again, by analyzing the options of "doing nothing" and licensing. In the modular system, the value of doing nothing equals the expected value of each module times the number of modules plus the incremental value of the system. In the proof to Proposition 4 we show that, for $0 < t < 1$, the payoff $\Pi_{\text{nothing, mod}}$ is strictly greater than Π_{nothing} and increasing in both the number of modules, *M,* and the degree of complementarity, β.

If the principal licenses all modules, he will get αV in total for the license and his own establishment will be worth the same. Thus the aggregate payoff to licensing is not changed by modularization.

Relational contracts

[10] For simplicity, we again assume modules are symmetric. Assuming different stand-alone values would simply add another parameter to each term in the summation, but there is no insight to be gained from the added complexity.

We next analyze the value of setting up relational contracts in a modular system. To calculate this value, we first determine the expected defection reward per agent within each module, assuming that defectors split the system reward (if any) evenly between modules. The agents within a module are in the same prisoners' dilemma or "chicken" situation discussed earlier. In contrast, defectors across modules are *not* in a prisoner's dilemma game: each one hopes that agents in other modules will defect. Thus, the necessary payment to keep each agent in a module loyal is:

$$\zeta V = \text{Expected Defection Reward per agent} \; = \; [(1-\beta)+p\beta]\,\alpha V/M. \tag{8}$$

where p denotes the probability that there is at least one defector in every module group.[11] This payment is strictly decreasing in the number of modules, and, if $p < 1$, in the system's degree of complementarity. We summarize our results in the following Proposition.

> ***Proposition 4.*** *When the system is symmetrically split into M modules, with β measuring the degree of complementarity between them, then if the share of trustworthy agents is positive and less than one (0 < t < 1), the values of the principal's various options change as follows.*
>
> ***(a)*** *The values of doing nothing and of paying agents increase strictly in M, the number of modules.*
> ***(b)*** *These values also increase strictly in β, the degree of complementarity.*
> ***(c)*** *The value of licensing remains unchanged.*

Proposition 4 applies within the bounds $0 < t < 1$. If $t = 1$ (everyone is trustworthy), then the strategy of doing nothing is trivially optimal and its value is invariant to both modularity and complementarity. If $t = 0$ (no one is trustworthy), then licensing is better than doing nothing, and is also invariant to modularity and complementarity. However, the strategy of agent payments may dominate licensing in this case. (If t

[11] By definition, p is the probability that there is at least one untrustworthy agent in each module. If the expected defection reward per module exceeds agent payments, then at least one untrustworthy agent per module will defect (if there is one) and p equals the probability that there is at least one defector in each module. If agent payments exceed the module-level reward, $(1-\beta)\,\alpha V/M$, but not the total expected defection reward per module, then agents across modules are in a coordination game: it pays to defect (for one agent per module at least) if, and only if, one or more agents from each other module defect. We assume that the principal expects agents to coordinate successfully, in which case, again, p equals the probability that at least one agent per module will defect.

= 0, the value of the agent payment strategy is strictly increasing in the number of modules and invariant to complementarity.)

Thus, in parallel with clans, when modularity is introduced, Region L in the (t,α) parameter space shrinks while the other two regions (N and P) expand. If, after modularization, the (t,α) combination lies in Regions N or P, then modularity unambiguously improves the outcome for the principal. Furthermore, as long as some agents are trustworthy, the effect is larger the greater the level of complementarity between modules.

We note that in a modular system, the principal could also apply hybrid strategies, i.e., treat individual modules differently. Analyzing such strategies in full generality is a rather complex exercise, which we omit in the interest of simplicity. For specific cases, we can show that hybrid strategies are inferior to one of the non-hybrid strategies, and the argument provides some indication that a hybrid strategy may never be optimal.[12]

Modularity with Clans

It is possible to have clans within a modular system. Recall that by definition, all members of a clan do the same thing: they either defect or they stay with the principal. We also assume that the principal knows the clan structure of his agents and can assign clan members to modules based on this knowledge. We assume there are L clans and the number of modules M is controlled by the principal. To begin with, for simplicity, we also assume that $L = M$, i.e., clan size and module size are matched. (This situation may arise, for example, when a team develops characteristics of a clan through working closely together on one module.)

[12] A specific hybrid strategy is for the principal to pay agents in Module 1 not to defect, and do nothing for all other modules. For strong complementarity (β close to 1), Module 1 becomes essential for profitable defection (so do all other modules, but those agents do not lose agent payments when they defect). Thus, even if there are defectors in the other modules this would only cause a minor loss (proportional to 1-β) to the principal if he can keep the agents in Module 1 loyal. However, by the same token these agents will be able to negotiate a large share of the system-level defection reward. This effect vitiates—partly or even entirely, depending on the precise assumptions—the principal's gains from having to pay only N/M instead of N agents. Under specific assumptions, one can show that paying agents in Module 1 only is inferior either to doing nothing or to symmetric agent payments for β=0, and inferior to the latter for β=1.

It is straightforward to show that the principal benefits from having all members of the same clan work on the same module. Clans, we have said, effectively reduce the number of independent decision-makers, which increases the value of the do–nothing strategy. They also reduce the per-person defection reward, thus increasing the value of the agent payment strategy. And they do not change the value of licensing. Hence, by the same reasoning used above, mapping clans onto modules increases the principal's payoff relative to modularity or clans alone in Regions N and P, and is value-neutral in Region L (where modularity and clans have no value anyway).

In contrast, if members of a clan are dispersed through the system, then an untrustworthy clan can be sure there is a willing defector in every module where one of its members works. The per-person payment needed to deter the clan's defection goes up, causing the value of the agent payment strategy to go down. The value of doing nothing also declines compared to a situation without clans because the principal will face system-level competition with higher probability.

Hybrid distributions of clan and non-clan members within and across modules change agents' expected rewards and the value of doing nothing in complicated, but computable ways. Nevertheless, concentrating clan members within modules is always better for both the agent payment strategy and the do-nothing strategy than spreading clan members across the system. In effect, a dispersed clan de-modularizes the system to some degree, because clan members can pool their knowledge about different modules and, if they defect, can recreate several modules or even the whole system.

EXTENSIONS

Our model can be extended in various respects. In this section, we discuss screening of agents and signaling by agents, legal protection of IP, imitation and substitution by third parties, value-increasing modularity, and fairness. We refrain from modeling these extensions formally in order to limit the complexity of the analysis, and instead provide a qualitative discussion of each issue.

Screening and signaling

An agent's prior exchange relationships may allow the principal to distinguish between trustworthy and untrustworthy agents, and may enable the agent to signal her type to the principal. If agents are otherwise identical and there is no scarcity of trustworthy agents in the labor market, then the principal will hire only those identified as trustworthy. If the selection mechanism works perfectly, then the problem is solved. However, this is unlikely. In identifying trustworthy agents there will generally be false positives (Granovetter, 1985). In that case, screening and signaling will serve to increase the share of trustworthy agents among the principal's employees above t, the value in the population, though not to 100 percent. Effectively, the t relevant to the principal's decision increases. As a result, the relevant (t,α) combination moves to the right in Figure 1, and the principal's payoff increases with further increases in t as soon as Region N ("do-nothing") is reached.

In a modular system, increasing t improves not only the value of doing nothing, but also that of paying agents. In this domain, the trustworthiness of the population interacts with system complementarity in an interesting way. If all agents are untrustworthy ($t = 0$), then there is no difference between complementary modules and strictly additive modules. In contrast, if some agents are trustworthy ($t > 0$), then required agent payments decline as complementarity increases. A more trustworthy population decreases the probability that an agent from every module will defect, which in turn increases the probability that the "system value" will be captured by the principal. In effect, in Regions N and P, t and β are strict complements as defined by Milgrom and Roberts (1990): actions that increase t or β make actions that increase the other variable more valuable.[13]

The situation is more complex if the principal has reason to employ known untrustworthy agents (e.g., because trustworthy agents are scarce, or because he could identify them after hiring but is unable to

[13] Within regions N and P, this fact can be demonstrated by simple calculus. If the principal switches regions, it follows from the fact that strict complementarity holds for the second-best strategy, hence must *a fortiori* hold for the first-best strategy.

lay them off), or if a superior screening technology obtains higher t for a new cohort of employees. In that case, rational choice theory may prescribe keeping the known untrustworthy agents (or the older cohort of employees, respectively) loyal through payments while doing nothing for the others. Again, this situation is perfectly amenable to analysis by our model. However, if agents learn about the differential treatment they may perceive it as unfair, an issue we discuss below.

The premise that trustworthy agents can be identified due to their prior exchange relationships has the additional implication that an agent's behavior with the focal principal will affect her reputation, and hence her ability to achieve attractive wages in the future (e.g., Hannah, 2005). We can account for this effect in our model by reducing the payoff to defectors by a reputational penalty. As a result, the agent payment strategy becomes less costly, hence more attractive to the principal, while doing nothing and licensing remain unaffected. Thus, Region P in the parameter space expands while the other regions contract.

Legal protection of intellectual property

A perfect legal system would enable the principal to obtain and enforce intellectual property (IP) rights and contracts. However, although worldwide IP rights have been strengthened by the recent TRIPS agreement, they are still weakly enforced in many emerging economies (Kyle and McGahan, 2009; Branstetter et al., 2011). And even in developed economies, there is generally some uncertainty about the enforceability and scope of IP protection (Lemley and Shapiro, 2007). Furthermore, legal systems are costly to use. Thus, in both emerging and developed economies, payments under relational contracts and modularity can supplement state-sanctioned IP rights.

In our model, legal protection of intellectual property would reduce the payoff to defection because of the possibility of legal sanctions such as fines or imprisonment.[14] Any reduction in defection rewards

[14] In the U.S. theft of trade secrets is the only violation of intellectual property law that carries potential criminal sanctions. (Lemley, 2008).

makes the agent payment strategy more attractive, and a reduction that turns the reward negative makes "doing nothing" the best option. Laws that brand the expropriation of organizational secrets as theft (and thus a violation of social norms) can also increase t, making the do-nothing strategy more attractive. However, the value of licensing remains unchanged. Thus, as with the introduction of clans and modularity, a legal system that protects intellectual property expands Regions N and P at the expense of Region L.

In the limit, with a perfect legal system, no agent perceives any benefit to defection and the do-nothing strategy prevails everywhere in the parameter space. Licensing, clans, relational contracts, and modularity are all irrelevant in this (admittedly unrealistic) world.

Imitation or substitution by third parties

Imitation or substitution by third parties may also threaten the principal's monopoly. If their identity is unknown, the principal cannot include such parties in any relational contract. But if imitation or substitution by external agents is likely, the value of the monopoly will decrease. While (rational) defection happens in the first stage of the repeated game or never, imitation or substitution occurs with some probability in every time period. And a legal system will protect the principal to some extent against imitation, but typically not against substitution (since IP rights generally cover the technical solution rather than the purpose of the system).

Extending our model to address the possibilities of imitation and substitution leaves the value of licensing unchanged, since by assumption only two firms can profitably operate in the market. In contrast, the values of doing nothing and of paying agents go down, since the principal's monopoly faces an additional (even if probabilistic) threat. Thus Region L expands, while Regions N and P shrink.

Modularity plausibly increases the risk of imitation or substitution of individual modules (Rivkin, 2000; Pil and Cohen, 2006; Ethiraj et al., 2008), an effect which has to be weighed against the protective effects of modularity against defection by agents. However, if would-be imitators or substitutors cannot attach their modules to the principal's system, then they can only appropriate the stand-alone value, under competition, of the respective module. Thus, an increase in the likelihood of module-level imitation or substitution is less detrimental for the principal the higher the degree of complementarity, β. In effect, as long as the probability that at least one module remains under the principal's exclusive control is positive, a higher degree of complementarity preserves more of the system value for the principal, thereby increasing the total value of the monopoly. We thus arrive at the important distinction between modularity in use, which increases the risk of imitation or substitution, and modularity in production, which mitigates the risk of expropriation by agents.

Value-increasing modularity

Modularization may increase system value and/or cause cost, two effects we have ignored so far. A value increase would affect both the value of monopoly and that of defection, and so could be addressed in our model, quite simply, by scaling the overall value, V. Similarly, a cost of modularization could be accounted for by a fixed cost term. Both changes would leave the mechanics and results of our model qualitatively unaffected.

Notice, however, that if the value of the monopoly, V, increases, then the dollar value of defection rewards increases as well. To counterbalance the higher rewards, agent payments must go up in absolute terms. In other words, a value-increasing modularization can have the effect of disequilibrating pre-existing relational contracts, unless the principal adjusts agent payments to reflect the new, higher value of the system. We will return to this point in our discussion of System/360 below.

Fairness

In designing a relational contract, rational choice theory recommends paying different groups of employees differently if they differ in terms of their share of trustworthy agents or the value of the module they are working on. The same recommendation arises if the principal employs clans of different sizes or a mixture of clans and individual agents. In our model, members of larger clans would be paid less than those in smaller clans, and any clan member would be paid less than an individual.

However, agent behavior may not be fully consistent with rational choice theory. This is particularly so for clans which are, by definition, groups of individuals who are socialized to obey the clan's norms (Ouchi, 1980, p. 132). Two very common norms, both borne out in laboratory experiments, are the norms of fairness and reciprocity. For example when humans play an "ultimatum" game against other humans, second movers regularly reject offers that they deem unfair. Players will punish someone who is unfair to them, even if it is against their immediate interest to do so, a behavior known as negative reciprocity (Prasnikar and Roth, 1992; Gächter and Fehr, 2002).

If agents' efforts and skill levels are the same, then paying some agents more than others is patently unfair. Thus, even though it may be rational and consistent with the Nash equilibrium to set up differential payments to prevent defection, the principal takes risks by doing so. Perceived unfairness can set in motion retaliatory strategies that are not individually or collectively rational. Individuals or clans might defect even though it is to their own loss.

That said, agents are not entirely irrational either. They will accept a certain level of perceived unfairness if it is in their own or their clan's interest. The principal can also take care that differential payments remain confidential (a common policy in many firms), or can justify them via organizational boundaries or nominally different job assignments.

EXAMPLES

In this section, we offer examples from practice that may help to clarify the assumptions, results, and limitations of our model. We begin with cases in which relational contracts and/or clans have been used to encourage loyalty and thus protect organizational secrets. We then discuss cases in which modularity, in conjunction with complementarity, has been used for this purpose. Our last example describes a value-creating modularization that upset the pre-existing relational contract between company and employees and triggered a large number of employee defections.

Relational Contracts, Clans, Screening and Signaling

Du Pont

In the United States in the 19th Century, the law regarding trade secrets protected documents and equipment, but not the knowledge in the heads of departing employees. Some employers did not hesitate to use relational contracts to ensure loyalty. For example, Massachusetts mill owner, Samuel Slater, paid his key employees higher wages to prevent their "aiding and assisting another mill" (Fisk, 2001, p. 467).

Not all mill owners were eager to use this practice, however. For example, in the early 19th Century, Irenee Du Pont founded an eponymous company to make gunpowder using secret formulas. The Du Ponts guarded their gunpowder secrets well. Through much of the 19th Century, all research was conducted by senior family members assisted by their sons and nephews (Chandler and Salsbury, 1971). However, workers at Du Pont mills also had valuable knowledge about powder-making processes and thus other mill owners sometimes tried to lure them away with offers of higher wages (Fisk, 2001). Notably Du Pont did not try to match these outside offers. He was very conscious of the high cost of paying all workers their defection reward, writing to another mill owner, "More than twenty other hands who … possess as much information as the ones you wish to bribe must naturally suppose they ought to receive the same exorbitant wages" (*ibid.*, p. 475).

In terms of our model, Du Pont preferred "doing nothing" to setting up a relational contract with his workers. However, there is indirect evidence that he relied in part on the embeddedness of his workers in kinship groups and communities that may have functioned as clans. Du Pont family members and their workers lived and worked side-by-side in relatively remote communities.[15] A defector from the company perforce would have to leave his home, friends and extended family behind. He would suffer the condemnation not only of the Du Ponts, but of his fellow workers who remained loyal to the firm.

---

[15] One does not locate gunpowder factories in the middle of cities.

Liftouts

Within organizations, people who work closely together may develop close social ties. For example, it is not uncommon for individuals to stay in an organization despite financially attractive outside offers because of their strong sense of loyalty towards colleagues, co-workers, and bosses. It is also possible for a group of employees with close ties to leave as one body. For example, when John Merriwether left Salomon Brothers, six other managing directors followed, to become partners in Merriwether's new firm, LTCM (Lowenstein, 2000).

The hiring away of an organizational unit or team is known as a "liftout." According to Groysberg and Abrahams (2006), liftouts are common in knowledge-based service businesses. A team that already knows how to work together can deliver better performance sooner than an equally skilled group of unconnected individuals. Such teams generally have common beliefs and mutually supportive social interactions. Whether they stay or leave the company, they act "as one" for the benefit of the group. Thus they conform to Ouchi's definition of a clan.

Modularity

We turn now to cases where modularity, in conjunction with complementarity, has been used to protect intellectual property.

Porcelain

In the introduction of this paper we described how Elector Frederick Augustus of Saxony used modularity to maintain a monopoly on European porcelain.[16] Augustus had to rely on one and later two agents—the inventor and his successors, respectively—to direct the porcelain-making process. Ironically, he could not rely on law—his law—to enforce intellectual property rights: a defector had only to ride as

[16] Augustus' motivations in owning a porcelain factory were complex. As a monarch, he maintained a large personal collection of porcelain objects and took the best and most ambitious pieces for himself. He also used the factory as a source of revenue. Both as a collector and as the sole owner of the factory, he was eager to maintain a monopoly over the porcelain-making process in Europe (Gleeson, 1998). Today his collection is located and can be viewed at the Zwinger Palace in Dresden. http://www.skd.museum/en/museums-institutions/zwinger-with-semperbau/porzellansammlung/ (viewed 12/24/13).

far as the nearest border (a relatively short distance) to escape his jurisdiction. Initially Augustus managed to keep all the essential knowledge in the head of one man ($N = 1$) whose movements he could control by force. Subsequently he split the knowledge of porcelain paste and glaze between two individuals. In doing so, Augustus modularized knowledge about the porcelain-making process. Applying Equation (8) above, note that the move from one agent and one module, $N = M = 1$, to two agents and two modules $N = M = 2$ potentially reduced total payments under the relational contract because the two modules were highly complementary: glazed porcelain products were much more valuable than either unglazed porcelain or glazed pottery. (The value of doing nothing and of licensing were unchanged.)

Radial Tires

Moving into the 20th Century, Liebeskind (1997) describes a similar split of production processes and knowledge related to making radial tires:

> During the 1960s, Michelin had a monopoly on knowledge relating to the production of high quality steel-belted radial tire manufacturing. In order to preserve this monopoly, manufacturing was divided into two separate processes: steel belt manufacturing, and tire production. Employees were not rotated between these manufacturing processes in a deliberate effort to restrict the number of employees that had knowledge about both processes. As a result, only a handful of very senior managers within Michelin were knowledgeable about the entire manufacturing process. (p. 645).

Practices in Emerging Economies

Emerging economies are an interesting source of examples for us because intellectual property rights are generally not well protected in these settings. As a result, the actions managers take to protect organizational secrets are more stark and visible.

Based on 120 interviews conducted in Brazil and Mexico in the 1980s, Sherwood (1990) reports that the following tactics were used to discourage the “predatory hiring” by competitors of employees with valuable knowledge: (1) Access to corporate technology was limited to family members or trusted employees. (2) Attractive housing was offered to key technical employees. (3) Critical technologies were worked on only by expatriate employees who had long-term career paths with the international parent firm. (4) New hires were exposed to only a small part of the overall operation and left in that role for several years, until they were viewed as trustworthy. (5) The founder alone knew the whole process, but a few life-long employees were permitted to know discrete parts. Note that practice (1) makes use of clans; (2) and (3) are forms of relational contracting; and (4) and (5) combine screening with modularity.

In a large sample study, Zhao (2006) looked at how modularity and complementarity were used to protect the value of multinational companies' R&D across international boundaries. To protect the value of their R&D, she argues, multinationals are likely to assign projects with little stand-alone value to R&D units in countries with weak IP protection. She goes on to present evidence from citation patterns that patents obtained by multinational subsidiaries in countries with weak IP rights have more value in conjunction with patents owned by the parent company than with patents owned by third parties.

Relatedly, in a series of case studies and interviews, Quan and Chesbrough (2010) found that multinational managers located projects with little stand-alone value in China because of concerns about weak IP protection in that country. The fact that the projects had little stand-alone value reduced defection rewards, hence the salaries needed to keep employees loyal. The multinationals could thus take advantage of the lower cost of conducting research in China without compromising the returns on their R&D investments.

IBM System/360

In general, our theoretical results indicate that modularity can be used to reduce the cost and/or risk of agents' expropriating valuable IP. How can these results be reconciled with the example of System/360, cited in the introduction, where modularization appeared to trigger a large number of employee defections with concomitant loss of IP? The answer to this conundrum is twofold.

First, the example does not fulfill our assumption that the principal only sells whole systems, or keeps third parties in other ways from attaching their modules to his system. Rather, to provide customers with configuration options, IBM sold its modules separately, providing modularity in use. This policy enabled its customers to integrate modules acquired from third parties into IBM systems. Thus, imitators could compete at the module level, without offering whole systems: the configurable system offered points of attachment for such third-party modules (Fixson and Park, 2008). Furthermore, IBM could be counted on to maintain a price umbrella over all System/360 modules: it was not going to slash prices to drive out small competitors (Christensen, 1993). In contrast, Augustus of Saxony did not sell unglazed porcelain, nor did he offer the service to glaze porcelain made by others. Also in the examples of Michelin and R&D in emerging economies, the principal did not offer individual modules.

Second, modularization can sometimes affect a system's value in ways that are unrelated to the protection of IP, but rather emerge from the option value associated with mixing and matching modules (Matutes and Regibeau, 1988; Baldwin and Clark, 2000). As the first "truly modular" computer (Ferguson and Morris, 1993), System/360 offered this possibility. Largely because of the options it gave to customers, System/360 was also a huge market success (Baldwin and Clark, 2000). A very high V meant that payments to knowledgeable agents necessary to prevent defection would have increased dramatically after System/360 was introduced and its success was apparent.

Defections from IBM started a few months after the launch of System/360. Interestingly, the first defectors had the characteristics of a clan: twelve individuals from the San Jose research lab left as a group to form a new company. Their former colleagues dubbed them the "dirty dozen." In the next five years, a significant number of IBM engineers, including some of the most creative and influential, left the company to join firms that were in direct competition with IBM (Pugh et al., 1991).

These defections can be understood as a response to the disequilibrium caused by a value-increasing modularization of very large proportions. As indicated, System/360's value as a product line was far greater than any of IBM's previous systems.[17] The disk and tape drive modules were only a subset of the system, but the cost of entering these markets was relatively low. Seventeen firms quickly entered the new market for "plug-compatible peripherals" (Transamerica vs. IBM, 698 F.2d 1377). Thus suddenly,

---

[17] In 1964, before System/360 was announced, IBM sold six separate computer systems. The average value tied up in each system was about $18 billion in today's dollars. At the end of 1967, the company's entire market value ($240 billion in today's dollars) was attributable to System/360 alone. (Baldwin and Clark, 2000.)

"people with knowledge of IBM technology and business plans [were] worth more outside the company … than inside" (Pugh et. al., 1991, p. 491). The result was "defections en masse" (*ibid.* p. 490).

Could IBM have prevented these defections? According to our model, it could have changed the relational contract to match the (perceived) defection rewards. However, such actions were anathema to IBM's senior executives at the time: they did not think in terms of defection rewards and agent payments, but in terms of trust and loyalty to the company. In the short run, IBM's managers elected to do nothing, and simply let the defectors go. It appears plausible that anticipatory licensing or an appropriate increase in loyalty payments, in line with our model, would have been advantageous to IBM.

CONCLUSION

A principal who derives rents from exclusive knowledge faces the threat of expropriation by agents. In this paper, we investigated the impact of modularity on intellectual property protection by formally modeling the threat of expropriation by agents. In our model, the principal has three options to address this threat: doing nothing, licensing the focal IP ex ante, and paying agents (via a relational contract) to prevent their defection. We showed that the principal can influence the value of these options by modularizing the technical system and by hiring clans of agents, thus exploiting relationships among them. His optimal choice depends on a number of external parameters—the percentage of trustworthy agents in the population, the intensity of competition in duopoly, and the degree of complementarity in the system. Extensions of the model can be used to understand the effects of screening and signaling in the hiring process, legal protection of intellectual property, imitation and substitution, disequilibrating changes in the value of knowledge, and social norms of fairness. We also presented examples to show how managers arrive at a strategy in practice.

We contribute to the theory of profiting from innovation in three ways. First, we show how the innovator's best choice of action against expropriation by agents—doing nothing, licensing, or paying agents—derives from the characteristics of the focal system, i.e., the number of agents, the share of trustworthy agents, the intensity of competition, the size of clans, the number of modules, and the degree of complementarity. We go beyond earlier work (Rønde, 2001; Rajan and Zingales, 2001) in considering a mixed population of trustworthy and untrustworthy agents. Heterogeneity in individual trustworthiness makes "doing nothing" a potentially attractive option and our model allows us to predict when it will be optimal. Furthermore, it links our analysis to the literature on screening and signaling, with the finding that imperfect screening has an effect only if the principal's best option, with screening, is to do nothing. Second, we extend existing work by showing how the innovator can use clans and modularity to increase his profits. We thus qualify the finding of earlier authors that modularity poses a risk to value appropriation (Rivkin, 2000; Pil and Cohen, 2006; Ethiraj, Levinthal, and Roy, 2008). Importantly, we carve out the distinction between modularity in use, which facilitates imitation and substitution, and modularity in production, which protects against expropriation by agents. If a system designer manages to achieve the latter while preventing the former, or can protect system interfaces from being used by third parties, then modularity should work to enhance value appropriation. We also show how clans and the modular architecture of the system interact to either reinforce or cancel each other. This interaction implies that an analysis of modularity alone will be misleading if clans are present in the focal organization. Third, we show how social relationships and norms of fairness affect the normative implications of an analysis based on rational choice theory. Since perceived fairness impacts an agent's behavior particularly strongly when trustworthiness matters, some of the recommendations of rational choice theory would yield the opposite of the intended outcome if not seen in the relevant social context.

Although the details determining the best strategies are complex, the implications for managers are relatively straightforward. The fundamental choices are (1) to protect the knowledge or not; and (2) to trust the agents or not. Relational contracts, that is, paying selected agents not to defect, makes it possible to protect knowledge and maintain a monopoly when agents are relatively untrustworthy. Clans, modularity, complementarity, and a legal system all serve to lower the cost and increase the value of this strategy. Trusting one's agents—what we have called "doing nothing"—is the most valuable course of action if it works, but is a risky strategy because trust can always be betrayed. Better screening and signaling technologies make it easier for the principal to trust his agents, but some residual risk always remains.

Our model has a number of limitations. Most importantly, we have presented agent payments and trust as stark alternatives. However, given norms of fairness and reciprocity, the boundary between these strategies tends to blur. Specifically, trustworthy agents may expect "fair" treatment from the principal, where "fair" entails some sharing of the value of the enterprise. In effect, t may be an implicit function of the agent payment parameter ζ. Then in an abstract sense, the principal's problem will be to determine a feasible and effective combination of t and ζ. It is certainly possible to set up a model of this type, but the $t(\zeta)$ function, if it exists, is not one we know much about.

Lastly, there are three potential routes to testing the model. The first is to conduct surveys and interviews, as in Sherwood (1990) and Poppo and Zenger (2002). These can determine whether some of the basic correlations predicted by the model, for example, a switch from agent payments to doing nothing as the perceived trustworthiness of employees increases, are observed in cross-section. However, such tests will be hampered by the fact that there is no guaranty of consistency in the perceptions of managers in different firms and countries. A second approach is to conduct case studies of events such as the introduction of System/360. However such events are rare and generally subject to multiple causal explanations. Finally, laboratory experiments can be used to test whether differences in t, α, or N lead to strategy choices consistent with the model's predictions. In other words, using intuitive reasoning alone, do individuals make choices that are consistent with the predictions of the model? And what, if any, role do norms of fairness and reciprocity play in determining their choices? The most promising route we think involves a combination of surveys, interviews and lab experiments to determine how managers reason practically about protecting organizational secrets.

ACKNOWLEDGEMENTS

We are grateful to Oliver Alexy, Juan Alcacer, Ron Adner, Ramon Casadesus-Masanell, Susan Cohen, Robert Gibbons, Venkat Kuppuswamy, Matt Marx, Christina Raasch, Andrew Torrance, Eric van den Steen, Eric von Hippel, Jason Woodard, Dennis Yao, and Minyuan Zhao, as well as participants in seminars and workshops at the Tuck School of Business, Harvard Business School, MIT, and UNC Kenan-Flagler Business School for comments that led to significant improvements of this paper. Errors and omissions are ours alone.

REFERENCES

Alvesson M, Lindkvist L. 1993. Transaction costs, clans and corporate culture. *Journal of Management Studies* **30**(3): 427-452.

Anton JJ, Yao DA. 2005. Markets for partially contractible knowledge: Bootstrapping versus bundling. *Journal of the European Economic Association* **3**(2/3): 745-754.

Arrow KJ. 1962. Economics of welfare and the allocation of resources for invention. In *The Rate and Direction of Inventive Activity*, Nelson, RR. (ed). Princeton, NJ: Princeton University Press; 609-625.

Baker G, Gibbons R, Murphy KJ. 2002. Relational contracts and the theory of the firm. *Quarterly Journal of Economics* **117**:39-84.

Baldwin CY, Clark KB. 2000. *Design Rules, Volume 1: The Power of Modularity.* MIT Press: Cambridge, MA.

Baldwin CY. 1983. Productivity and labor unions: An application of the theory of self-enforcing contracts. *The Journal of Business* **56**(2):155-185.

Barney JB. 1991. Firm resources and sustained competitive advantage. *Journal of Management* **17**: 99–120.

Bjørnskov C. 2007. Determinants of generalized trust: A cross-country comparison. *Public Choice* **130**(1-2): 1-21.

Branstetter LG, Fisman R, Foley CF, Saggi K. 2011. Does intellectual property reform spur industrial development. *Journal of International Economics* **83**(2011): 27-36.

Brynjolfsson E. 1994. Information assets, technology and organization. *Management Science* **40**(12): 1645-1662.

Bull C. 1987. The existence of self-enforcing implicit contracts. *Quarterly Journal of Economics* **102**(1): 147-159.

Chandler A D, Salsbury S. 1971. *Pierre S. DuPont and the Making of the Modern Corporation*. Harper Row: New York.

Christensen CM. 1993. The rigid disk drive industry: a history of commercial and technological turbulence. *Business History Review* **67**(4): 531-588.

Coase RH. 1937. The nature of the firm. *Economica* **4**(16): 386-405.

Dyer JH, Singh H. 1998. The relational view: Cooperative strategy and sources of interorganizational competitive advantage. *Academy of Management Review* **23**(4): 660-679.

Ethiraj SK, Levinthal D, Roy R. 2008. The Dual Role of Modularity: Innovation and Imitation. *Management Science* **54**(5): 939-955.

Fixson SK, Park J-K. 2008. The Power of Integrality: Linkages between Product Architecture, Innovation and Industry Structure. *Research Policy* **37**(8): 1296-1316.

Ferguson CH, Morris CR. 1993. *Computer Wars: How the West Can Win in a Post-IBM World*. Times Books: New York.

Fisk C. 2001. Working Knowledge: Trade Secrets, Restrictive Covenants in Employment, and the Rise of Corporate Intellectual Property, 1800-1920. *Hastings Law Journal* **52**(2): 441-528.

Gächter S, Fehr E. 2002. Fairness in the Labour Market? A Survey of Experimental Results. Zurich IEER Working Paper No. 114. Available at SSRN: http://ssrn.com.ezp-prod1.hul.harvard.edu/abstract=310619.

Garud R, Kumaraswamy A. 1993. Changing competitive dynamics in network industries: An exploration of sun microsystems' open systems strategy. *Strategic Management Journal* **14**: 351-369.

Garud R, Kumaraswamy A. 1995. Technological and organizational designs for realizing economies of substitution. *Strategic Management Journal* **16**: 93-109.

Gibbons R, Henderson R. 2012. Relational contracts and organizational capabilities. *Organization Science* **23**: 1350-64.

Gleeson J. 1998. *The Arcanum: The Extraordinary True Story*. Warner Books: New York.

Granovetter M. 1985. Economic action and social structure: the problem of embeddedness. *American Journal of Sociology* **91**(3): 481-510.

Greif A. 1998. Self-enforcing political systems and economic growth: Late medieval Genoa. In *Analytic Narratives*, Bates et al (eds.). Princeton University Press: Princeton, NJ.

Grossman SJ, Hart OD. 1986. The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of Political Economy* **94**(4): 691-719.

Groysberg B, Abrahams R. 2006. Lift outs: how to acquire a high-functioning team. *Harvard Business Review* **84**(12): 133.

Hannah, DR. 2005. Should I keep a secret? The effects of trade secret protection procedures on employees' obligations to protect trade secrets. *Organization Science* **16**(1): 71-84.

Hart OD, Moore J. 1990. Property rights and the nature of the firm. *Journal of Political Economy* **98**(6): 1119-1158.

Johnson JP, Korsgaard MA, Sapienza HJ. 2002. Perceived fairness, decision control, and commitment in international joint venture management teams. *Strategic Management Journal* **23**: 1141–1160.

Kreps D. 1990. Corporate culture and economic theory. In *Perspectives on Positive Political Economy,* Alt J, Shepsle K. (eds). Cambridge University Press: 90-143.

Kyle M, McGahan A. 2009. Investments in Pharmaceuticals before and after TRIPS. Working paper 15468 (October). National Bureau of Economic Research.

Langlois RN, Robertson PL. 1992. Networks and innovation in a modular system: Lessons from the microcomputer and stereo component industries. *Research Policy* **21**: 297–313.

Lemley MA. 2008. The Surprising Virtues of Treating Trade Secrets as IP Rights, *Stanford Law Review* **61**: 311.

Lemley MA, Shapiro C. 2007. Patent holdup and royalty stacking. *Texas Law Review* **85**: 1991–2049.

Liebeskind JP. 1997. Keeping organizational secrets: Protective institutional mechanisms and their costs. *Industrial and Corporate Change* **6**: 623-663.

Lowenstein R. 2000. *When Genius Failed: The Rise and Fall of Long-Term Capital Management*. Random House, New York.

Macneil IR. 1978. Contracts: adjustment of long-term economic relations under classical, neoclassical, and relational contract law. *Northwestern University Law Review* **72**: 854-905.

Macneil IR. 1985. Relational contract: what we do and do not know. *Wisconsin Law Review* 483-525..

Macneil IR. 1987. Relational contract theory as sociology: a reply to Professors Lindenberg and de Vos. *Journal of Institutional and Theoretical Economics* **143**(2): 272-290.

Marx M. 2011. The firm strikes back: Non-compete agreements and the mobility of technical professionals. *American Sociological Review* **76**(5): 695-712.

Maskus KE. 2000. *International Property Rights in the Global Economy*. Institute for International Economics:Washington, D.C.

Matutes C, Regibeau P. 1988. “Mix and match”: Product compatibility without network externalities. *RAND Journal of Economics* **19**: 221-234.

Mazar N, Amir O, Ariely D. 2008. The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research* **45**(6): 633-644.

Mead C, Conway L. 1980. *Introduction to VLSI Systems*, Addison-Wesley, Reading, MA.

Milgrom P, Roberts J. 1990. The economics of manufacturing: Technology, strategy and organization. *American Economic Review* **80**(3): 511-28.

Ouchi WG. 1980. Markets, bureaucracies, and clans. *Administrative Science Quarterly* **25**: 129-141.

Oxley JE. 1997. Appropriability hazards and governance in strategic alliances: A transaction cost approach. *Journal of law, Economics, and Organization* **13**(2): 387-409.

Oxley JE. 1999. Institutional environment and the mechanisms of governance: the impact of intellectual property protection on the structure of inter-firm alliances. *Journal of Economic Behavior & Organization* **38**(3): 283-309.

Pil FK, Cohen SK. 2006. Modularity: Implications for imitation, innovation, and sustained competitive advantage. *Academy of Management Review* **31**(4): 995-1011.

Poppo L, Zenger T. 2002. Do formal contracts and relational governance function as substitutes or complements? *Strategic Management Journal* **23**(8): 707-725.

Poppo L, Zhou KZ, Zenger TR. 2008. Examining the conditional limits of relational governance: specialized assets, performance ambiguity, and long-standing ties. *Journal of Management Studies* **45**(7): 1195-1216.

Posner EA. 2002. *Law and Social Norms*. Harvard University Press.

Prasnikar V, Roth AE. 1992. Considerations of fairness and strategy: Experimental data from sequential games. *Quarterly Journal of Economics* **107**(3): 865-888.

Pugh EW, Johnson LR, Palmer JH. 1991. *IBM's 360 and Early 370 Systems.* MIT Press: Cambridge, MA.

Quan X, Chesbrough H. 2010. Hierarchical segmentation of R&D process and intellectual property protection: Evidence from multinational R&D laboratories in China. *IEEE Transactions in Engineering Management* **57**(1): 9-21.

Rajan R, Zingales L. 2001. The firm as a dedicated hierarchy: A theory of the origin and growth of firms. *Quarterly Journal of Economics* **116**(1): 805-851.

Rivkin JW. 2000. Imitation of Complex Strategies. *Management Science* **46**:824-844.

Rønde T. 2001. Trade secrets and information sharing. *Journal of Economics & Management Strategy* **10**(3), 391–417.

Rotter JB. 1980. Interpersonal trust, trustworthiness, and gullibility. *American Psychologist* **35**(1): 1.

Sako M. 1998. Does trust improve business performance? In *Organizational Trust: A Reader*, Kramer R. (ed.) Oxford University Press: Oxford, UK; 88-117.

Sanchez R, Mahoney JT. 1996. Modularity, flexibility, and knowledge management in product and organizational design. *Strategic Management Journal* Winter Special Issue **17**: 63–76.

Schilling MA. 2000. Toward a general modular systems theory and its application to interfirm product modularity. *Academy of Management Review* **25**: 312–334.

Sherwood RM. 1990. *Intellectual Property and Economic Development*. Boulder: Westview Press.

Simon HA. 1962. The architecture of complexity. *Proceedings of the American Philosophical Society* **106**: 467–482.

Teece DJ. 1986. Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy* **15**: 285–305.

Teece DJ. 2000. Strategies for managing knowledge assets: the role of firm structure and industrial context. *Long Range Planning* **33**(1): 35–54.

Telser LG. 1980. A Theory of Self-Enforcing Agreements. *Journal of Business* **53**(1): 27-44.

Ulrich KT, Eppinger SD. 1994. *Product Architecture, Methodologies for Product Design and Development*. McGraw-Hill: New York.

Uzzi B. 1997. Social structure and competition in interfirm networks: The paradox of embeddedness. *Administrative Science Quarterly* **42**(1): 35-67.

von Hippel E. 1990. Task partitioning: An innovation process variable. *Research Policy* **19**: 407-18.

Williamson OE. 1993. Calculativeness, trust, and economic organization. *Journal of Law and Economics* **36**: 453-486.

Yamagishi T, Yamagishi M. 1994. Trust and commitment in the United States and Japan. *Motivation and Emotion* **18**(2): 129-166.

Zhao M. 2006. Conducting R&D in countries with weak intellectual property rights protection. *Management Science* **52**: 1185–1199.

APPENDIX

Proof of Proposition 1

The proof of Proposition 1(a) follows from straightforward calculation. Since $1/N > 0$ and since $\alpha/(1-\alpha)$ increases in α, also $[\alpha/(1-\alpha)]^{1/N}$ increases in α. Since $\alpha < ½$ we have $\alpha/(1-\alpha) < 1$. Thus $[\alpha/(1-\alpha)]^{x}$ is decreasing in x (for any $x > 0$) and $[\alpha/(1-\alpha)]^{1/N}$ is increasing in N. This proves part (b) of the proposition.

Proof of Proposition 2

Part (a) follows from the argument in the main text. Part (b) is shown by equating the respective payoff terms (e.g., $\Pi_{\text{payments}} = \Pi_{\text{nothing}}$) and solving for N and t, respectively. For part (c), the derivatives are after some algebraic manipulation:

$$\frac{\partial t^*}{\partial \alpha} = \frac{-1}{(1-\alpha)^2}\left[1-\frac{\alpha N}{1-\alpha}\right]^{\frac{1}{N}-1}, \quad \frac{\partial t^*}{\partial N} = \frac{-1}{N^2}\left[1-\frac{\alpha N}{1-\alpha}\right]^{\frac{1}{N}}\left\{\frac{\alpha N}{1-\alpha(N+1)} - \ln\left(1+\frac{\alpha N}{1-\alpha(N+1)}\right)\right\}. \qquad \text{(A-1)}$$

Since t^* is only defined for $\mathrm{N} < \alpha^{-1} - 1$, the terms in square brackets are positive. It follows directly that the first derivative is negative. For the second derivative, note that $\ln(1+x) < x$ for $x > 0$ (since $\ln(1)=0$, $\ln'(1)=1$, and $\ln''(x) < 0$ for all x). It follows that the term in curly brackets is positive, and hence the derivative overall negative.

Proof of Proposition 3

We rewrite Equation 1 as:

$$\Pi_{\text{nothing}} = [\alpha + t^N(1-\alpha)]V \ . \qquad \text{(A-2)}$$

Given that $\alpha < 1$, if $0 < t < 1$ then $t^N(1-\alpha)$ decreases with N; otherwise it remains unchanged. Thus Π_{nothing} decreases in N (or remains unchanged if t=0 or t=1). For From Equation 3, it is clear that Π_{payment} decreases in N if $\alpha > 0$ and remains unchanged otherwise. $\Pi_{\text{licensing}}$, finally, does not depend on N.

Proof of Proposition 4

(a) We first consider how the values of doing nothing and of paying agents vary with *M*. We assume, as stipulated in the Proposition, that $0 < t < 1$.

Doing nothing: In the modular system, the value of doing nothing equals the expected value of each module times the number of modules plus the incremental value of the system. Thus we can write the value of the do-nothing strategy for the modular system as:

$$\Pi_{\text{nothing,mod}} = (1-\beta)\,[g(t,\alpha)] + \beta\,[h(t,\alpha)]\,, \tag{A-3}$$

where $g(t,\alpha)$ is the expected value of the *M* modules and $h(t,\alpha)$ is the corresponding expected system-level value.

Using Equation 1 in the text and the symmetry of modules, we have:

$$g(t,\alpha) = \left[t^{N/M} + \alpha\left(1 - t^{N/M}\right)\right] V \tag{A-4}$$

Note that $g(t,\alpha)$ increases in *M* (as long as $0 < t < 1$).

Turning to the second term in square brackets, we have:

$$h(t,\alpha) = [(1-p) + \alpha p]\, V\,, \tag{A-5}$$

where *p* denotes the probability that someone is untrustworthy *in every module*. The probability that at least one agent is untrustworthy in any module is $1 - t^{N/M}$, thus:

$$p = \left(1 - t^{N/M}\right)^{M}. \tag{A-5}$$

Note that $p = \left(1 - t^{N/M}\right)^{M}$ decreases in *M* in two respects: increasing *M* in the ratio N/M increases $t^{N/M}$ and thus decreases the overall expression; and increasing *M* in the external exponent decreases p because the term in brackets is between zero and one. Since *p* decreases in *M*, $h(t,\alpha)$ increases in *M* (again as long as $0 < t < 1$).

If $M > 1$, then $t^{N/M} > t^{N}$ and thus $\Pi_{\text{nothing}} < g(t,\alpha)$. Furthermore, by Equation A-5, $p < 1 - t^{N/M}$, thus $1 - p > t^{N/M}$ and $g(t,\alpha) < h(t,\alpha)$. Using the inequalities, we have:

$$\Pi_{\text{nothing}} \;<\; g(t,\alpha) \;<\; (1-\beta)g(t,\alpha) + \beta h(t,\alpha) \;=\; \Pi_{\text{nothing,mod}} \;. \tag{A-6}$$

Furthermore, we have seen that both $g(t,\alpha)$ and $h(t,\alpha)$ are increasing in *M*, thus, within the stipulated range, $\Pi_{\text{nothing,mod}}$ also increases with *M*.

Paying agents: From Equation 8 the necessary payment to keep each agent loyal is:

$$\text{Necessary payment per agent} = [(1-\beta) + p\beta]\;\alpha V/M \tag{A-7}$$

The necessary payment is less than αV if $M > 1$. Because *p* and *1/M* are both decreasing in *M*, the necessary payment is likewise decreasing in the number of modules. Thus modularization unambiguously increases the value of this strategy.

(b) We next consider how the values of doing nothing and of paying agents vary with β.

Doing Nothing: We can rewrite (A-3) as:

$$\Pi_{\text{nothing,mod}} = \; g(t,\alpha) + \beta[h(t,\alpha) - g(t,\alpha)] \;. \tag{A-8}$$

Since $g(t,\alpha) < h(t,\alpha)$, it follows immediately that $\Pi_{\text{nothing,mod}}$ increases with β.

Paying agents: It follows directly from (A-7) that, unless $p = 1$, the necessary payment is strictly decreasing in β.

***(c)* Licensing:** If the principal licenses all modules, he will get αV in total. Thus modularization does not change the value of licensing.

FIGURE

Figure 1. The principal's best option as a function of parameters t and α